













# TeraChem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics

Stefan Seritan<sup>1,2</sup>  | Christoph Bannwarth<sup>1,2</sup>  | Bryan S. Fales<sup>1,2</sup>  |  
 Edward G. Hohenstein<sup>1,2</sup>  | Christine M. Isborn<sup>3</sup> |  
 Sara I. L. Kokkila-Schumacher<sup>4</sup>  | Xin Li<sup>5</sup>  | Fang Liu<sup>6</sup>  |  
 Nathan Luehr<sup>7</sup>  | James W. Snyder Jr.<sup>8</sup> | Chenchen Song<sup>9,10</sup>  |  
 Alexey V. Titov<sup>11</sup> | Ivan S. Ufimtsev<sup>12</sup>  | Lee-Ping Wang<sup>13</sup>  |  
 Todd J. Martínez<sup>1,2</sup> 

<sup>1</sup>Department of Chemistry and the PULSE Institute, Stanford University, Stanford, California

<sup>2</sup>SLAC National Accelerator Laboratory, Menlo Park, California

<sup>3</sup>Department of Chemistry, University of California Merced, Merced, California

<sup>4</sup>IBM Thomas J. Watson Research Center, Yorktown Heights, New York

<sup>5</sup>Division of Theoretical Chemistry and Biology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm, Sweden

<sup>6</sup>Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

<sup>7</sup>NVIDIA, Santa Clara, California

<sup>8</sup>Adobe, San Jose, California

<sup>9</sup>Department of Physics, University of California Berkeley, Berkeley, California

<sup>10</sup>Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California

<sup>11</sup>Intel Corporation, Santa Clara, California

<sup>12</sup>Department of Structural Biology, Stanford University School of Medicine, Stanford, California

<sup>13</sup>Department of Chemistry, University of California Davis, Davis, California

## Correspondence

Todd J. Martínez, Department of Chemistry and the PULSE Institute, Stanford University, Stanford, CA 94305.  
 Email: toddjmartinez@gmail.com

## Funding information

Boeing; Deutsche Akademie der Naturforscher Leopoldina - Nationale Akademie der Wissenschaften, Grant/Award Number: LPDS 2018-09; Office of Science

## Abstract

TeraChem was born in 2008 with the goal of providing fast on-the-fly electronic structure calculations to facilitate ab initio molecular dynamics studies of large biochemical systems such as photoswitchable proteins and multichromophoric antenna complexes. Originally developed for videogaming applications, graphics processing units (GPUs) offered a low-cost parallel computer architecture that became more accessible for general-purpose GPU computing with the release of CUDA in 2007. The evaluation of the electron repulsion integrals (ERIs) is a major bottleneck in electronic structure codes and provides an attractive target for acceleration on GPUs. Thus, highly efficient routines for evaluation of and contractions between the ERIs and density matrices were implemented in TeraChem. Electronic structure methods were developed and implemented to leverage these integral contraction routines, resulting in the first quantum

chemistry package designed from the ground up for GPUs. This GPU acceleration makes TeraChem capable of performing large-scale ground and excited state calculations in the gas and condensed phase. Today, TeraChem's speed forms the basis for a suite of quantum chemistry applications, including optimization and dynamics of proteins, automated and interactive chemical discovery tools, and large-scale nonadiabatic dynamics simulations.

This article is categorized under:

Electronic Structure Theory > Ab Initio Electronic Structure Methods

Software > Quantum Chemistry

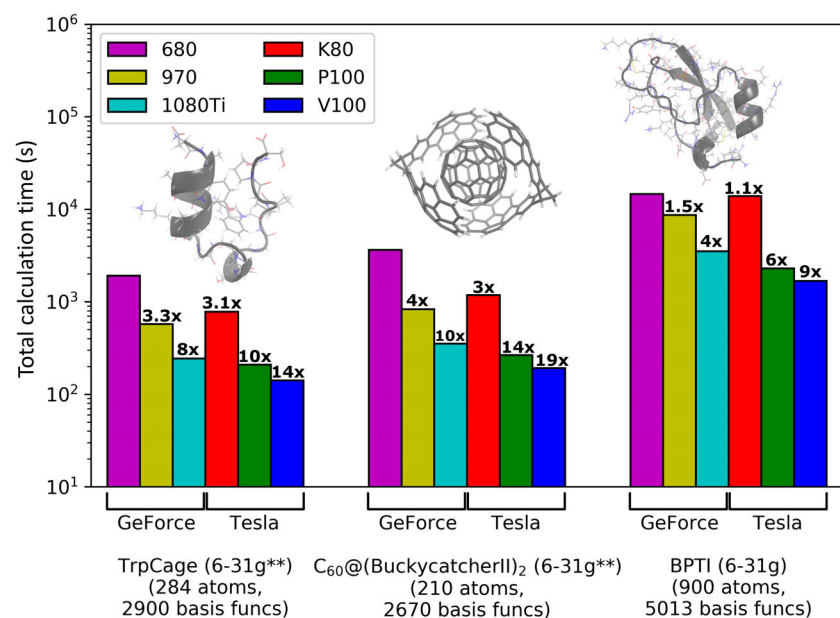
Structure and Mechanism > Computational Biochemistry and Biophysics

#### KEYWORDS

Electronic structure, ab initio molecular dynamics, graphical processing units

## 1 | HISTORICAL OVERVIEW

Graphics processing units (GPUs) are routinely used in consumer videogame hardware and can be viewed as highly efficient data-parallel computer architectures. We began exploring the use of GPUs for electronic structure and first principles dynamics calculations in the context of Sony PlayStation 2 consoles,<sup>1</sup> before modern scientific-focused discrete GPUs were available. Progress was hampered by the lack of a well-developed software ecosystem and the proprietary nature of the PlayStation 2. In 2007, NVIDIA released the Compute Unified Device Architecture<sup>2</sup> (CUDA) framework that greatly simplified the development of general-purpose GPU software. Thus, we rapidly switched our focus to CUDA-capable GPUs and the TeraChem package was born. Multiple order of magnitude performance advantages (compared to some existing CPU-based electronic structure codes) were evident already in the first reports of the GPU-based algorithms developed in TeraChem.<sup>3–6</sup> About one order of magnitude of this performance advantage could be explained by the computational power of the GPU (a performance differential that still exists today) and the remainder of these original gains came from the new data-parallel algorithms that were developed to use the unique streaming multiprocessor architecture of GPUs effectively. Fortuitously, these initial efforts occurred just as Moore's law was beginning to falter for CPUs. The majority of CPU performance improvements in recent years come from the introduction of optimized vectorization instruction sets (e.g., AVX2, AVX512) that require some algorithm redevelopment.



**FIGURE 1** Performance of single point energy and gradient calculations for TrpCage (PDB ID: 2JOF, left), a BuckycatcherII complex (center), and bovine pancreatic trypsin inhibitor (PDB ID: 6PTI, right) over several generations of GeForce and Tesla GPUs. The same Hartree–Fock implementation is used on all GPUs and speedups are reported compared to the GeForce GTX 680 for each molecule. Each calculation used a single GPU and single core of an Intel Xeon CPU, either a 2.8 GHz E5-2680 (K80), a 2.5 GHz E5-2640 CPU (680/970), a 2.6 GHz E5-2660v3 (P100), or a 3.4 GHz E5-2643v4 (1080Ti/V100)

Figure 1 shows that the performance of TeraChem continues to increase significantly without any additional code optimizations as each new GPU generation provides substantially more compute power and memory bandwidth.

GPUs reach maximum performance when working with single precision and when CPU-GPU memory transfers are minimized.<sup>3,7</sup> In fact, the earliest commodity GPUs aimed at the videogame market lacked hardware support for double precision arithmetic. Although this situation has changed and modern graphics cards can also handle double precision, they still perform better in single precision due to the lower memory requirements and optimized hardware pipelines. This limitation was not much of a hindrance for the application of GPUs to classical force-field molecular dynamics,<sup>8–12</sup> where single precision can be sufficient. However, quantum chemistry typically requires double precision to handle the large values for the total electronic energies and molecular orbital coefficients. Thus, balancing accuracy and performance in the face of precision limitations was an important consideration when developing GPU-based algorithms for electronic structure.

TeraChem expands the electronic wavefunction in a basis of contracted atomic orbitals  $\phi_\mu(\vec{r})$  which are linear combinations of atom-centered Cartesian Gaussian-type orbitals  $\chi_k(\vec{r})$ , often called primitive CGTOs (or just “primitives”):

$$\phi_\mu(\vec{r}) = \sum_k c_{\mu k} \chi_k(\vec{r}), \quad (1)$$

$$\chi_k(\vec{r}) = N(x - X_A)^{n_x} (y - Y_A)^{n_y} (z - Z_A)^{n_z} e^{-\zeta |\vec{r} - \vec{R}_A|^2}, \quad (2)$$

where  $N$  is a normalization factor,  $\vec{R}_A = [X_A, Y_A, Z_A]$  is the coordinate of the  $A$ th nucleus and the total angular momentum of the CGTO is given as  $L = n_x + n_y + n_z$ . The computational bottleneck in electronic structure codes is often the evaluation of the two-electron integral tensor, also known as the electron repulsion integrals (ERIs), whose size formally scales as  $O(N^4)$  with the number of orbitals  $N$ . It is neither efficient nor desirable to compute the ERIs in isolation; instead, various contractions of the ERIs with density matrices can be used directly. For example, Fock matrix construction requires the following two integral contractions:

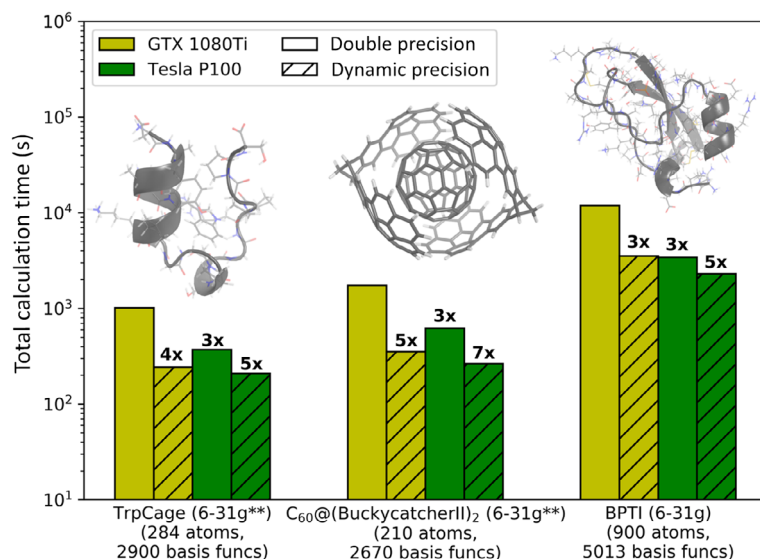
$$J_{\mu\nu} = \sum_{\lambda\sigma} (\mu\nu|\lambda\sigma) P_{\lambda\sigma}, \quad (3)$$

$$K_{\mu\nu} = \sum_{\lambda\sigma} (\mu\lambda|\nu\sigma) P_{\lambda\sigma}, \quad (4)$$

where  $(\mu\nu|\lambda\sigma)$  is an ERI over contracted atomic orbitals and  $P_{\lambda\sigma}$  is a density matrix. Typically, Equation (3) is known as the Coulomb matrix build (or “J-build”), and Equation (4) is known as the exchange matrix build (or “K-build”).

In 2008, Ufimtsev demonstrated the ability of GPUs to accelerate the direct evaluation of the Coulomb matrix,<sup>4</sup> exchange matrix and subsequent Fock matrix construction,<sup>5</sup> and analytic gradients for the self-consistent field (SCF) procedure,<sup>6</sup> utilizing a mixed precision scheme with all accumulations in double precision. It quickly became clear that the GPU architecture was well-suited to provide fast J and K builds. As at least partially foreseen by Almlof,<sup>13</sup> the formal  $O(N^4)$  scaling is reduced in practice to  $O(N^2)$  by exploiting element sparsity, caused by the inherent locality of the contracted AO basis, through presorting and screening the primitive AO pairs in the bra and ket of the primitive integrals. Later, Fock matrix diagonalization was also parallelized through the use of MAGMA, a library for GPU-accelerated linear algebra, for a fully GPU-accelerated SCF code.<sup>14</sup> PetaChem, LLC was founded in 2009 by Ufimtsev and Martínez to support the continued development of TeraChem based on these GPU-accelerated J and K builds, and the first commercial version of the software was released in May 2010.

A dynamic precision scheme was introduced to maintain double precision accuracy while maximally exploiting the speed of single precision arithmetic.<sup>7</sup> On commodity videogaming cards, this can lead to as much as a fivefold increase in performance, as shown in Figure 2. Even for scientific grade GPUs, this typically leads to a twofold performance increase (largely because of the increased bandwidth when dealing with single-precision compared to double-precision numbers). Accuracy of the J and K builds is controlled by two thresholds: integrals with a density-weighted Schwarz bound smaller than the first threshold are neglected, while the second threshold determines which integrals are treated in single or double precision in mixed/dynamic precision schemes. It can be helpful to think of this as three separate levels of precision—the largest integrals are computed in full double precision, medium-sized integrals are computed in single precision, and the smallest integrals are completely neglected (i.e., computed with no precision). With dynamic



**FIGURE 2** Comparison of the double and dynamic precision schemes<sup>7</sup> between Pascal-generation GeForce and Tesla GPUs. The same Hartree–Fock implementation is used on all GPUs and speedups are reported compared to double precision on the GeForce GTX 1080Ti for each molecule. Each calculation used a single GPU and single core of an Intel Xeon CPU as described in Figure 1

precision, the definitions of “large,” “medium-sized,” and “small” change during the course of the calculation to avoid wasteful computation of overly precise integrals.

Titov used code generation techniques to develop optimal GPU architecture-specific kernels through  $d$  angular momentum.<sup>15</sup> Access to optimized routines for these ERI contraction operations also began influencing the development and implementation of new post-SCF methods. Just as numerical linear algebra algorithms see performance gains when cast in terms of optimized matrix–vector and matrix–matrix operations, the J and K builds in the non-orthogonal Cartesian AO basis became two core electronic structure computational primitives due to their extremely efficient computation on the GPU. All subsequent electronic structure implementations in TeraChem use these highly optimized basic computational primitives by feeding generalized density matrices as inputs, therefore leveraging GPUs effectively through the entire code. For example, Isborn and Luehr added configuration interaction singles (CIS) and Tamm–Dancoff approximation time dependent density functional theory (TDA-TDDFT) by substituting nonsymmetric transition densities instead of the symmetric ground state densities shown in Equations (3) and (4).<sup>16,17</sup> These methods were later extended to the full TDDFT<sup>18</sup> (also known as the random phase approximation, or RPA). In 2014, Hohenstein developed the first fully AO-driven implementation of complete active space (CAS) methods including complete active space self-consistent field (CASSCF) and floating occupation molecular orbital complete active space configuration interaction (FOMO-CASCI).<sup>19–21</sup> These methods were leveraged in excited-state dynamics simulations of processes of several picoseconds including more than 600 quantum mechanical atoms.<sup>22</sup>

In 2015, Fales implemented a GPU-accelerated determinant-based direct configuration interaction (CI) program.<sup>23</sup> In determinant-based CI, the wavefunction is expressed as a linear combination of all possible determinants for a given number of active electrons and orbitals:

$$|\Psi\rangle = \sum_I c_I |\Phi_I\rangle. \quad (5)$$

Three electronic structure quantities were identified as core components in CI theory: generalized one- and two-particle density matrices (OPDMs/TPDMs) and Hamiltonian-CI vector products (also called  $\sigma$  builds).

$$\gamma_{pq} = \sum_{IJ} c_I c_J \langle \Phi_I | \hat{E}_{pq} | \Phi_J \rangle, \quad (6)$$

$$\Gamma_{pqrs} = \frac{1}{2} \sum_{IJ} c_I c_J \langle \Phi_I | \hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} | \Phi_J \rangle, \quad (7)$$

$$\sigma_I = \sum_J H_{IJ} c_J, \quad (8)$$

where  $\hat{E}_{pq}$  is an excitation operator from molecular orbital  $p$  to  $q$  and  $H_{IJ}$  is an element of the CI Hamiltonian matrix. Using the J, K, OPDM, TPDM, and  $\sigma$  building blocks, Snyder and Hohenstein developed an AO-driven implementation of the coupled-perturbed state-averaged CASSCF (CP-SA-CASSCF) equations, enabling the evaluation of analytic derivatives and nonadiabatic coupling elements with the state-averaged CASSCF (SA-CASSCF) method.<sup>24</sup> These developments allowed TeraChem to perform the largest nonadiabatic dynamics simulation with SA-CASSCF potential surfaces at the time of writing.<sup>25</sup> The GPU-accelerated CI program and the AO-driven CP-SA-CASSCF were later combined by Snyder and Fales to allow efficient SA-CASSCF calculations with large active spaces.<sup>26</sup>

Parallel developments in TeraChem include leveraging rank sparsity in the ERI tensor through the tensor hypercontraction (THC) framework.<sup>27–29</sup> Kokkila-Schumacher and Song applied GPU-accelerated THC to several perturbative methods, including second-order Møller-Plesset (MP2),<sup>30,31</sup> scaled-opposite-spin MP2 (SOS-MP2),<sup>31–33</sup> second-order approximate coupled-cluster singles and doubles (CC2),<sup>34,35</sup> and second-order CAS perturbation theory (CASPT2).<sup>36</sup> It is clear that the THC framework plays a similar role to the Coulomb and exchange matrix builds above; specifically, THC provides a reduced-scaling pathway for different types of integral contractions over the ERI tensor. Another parallel development is the introduction of two new potential electronic structure building blocks by Liu involving primitive AO pair-surface charge Coulomb interactions for a GPU-based implementation of the conductor-like polarizable continuum model (C-PCM).<sup>37</sup>

One decade after TeraChem's inception, these numerous developments have culminated in an electronic structure package that enables large-scale ab initio calculations on workstation-class hardware. Recent development effort has been placed in establishing well-defined interfaces and encapsulating the various GPU-accelerated computational primitives into standalone libraries. Additionally, a new language-agnostic socket-based interface allows TeraChem to run as a server for single-point calculations, providing access to fast electronic structure from high-level languages like Python and the ability to deploy TeraChem in flexible workflows on modern distributed computing resources such as commercial cloud platforms.<sup>38</sup>

## 2 | SCF AND DENSITY FUNCTIONAL THEORY METHODS

The Hartree–Fock (HF) or Kohn–Sham equations form the basis of all electronic structure packages; to that end, TeraChem provides efficient implementations of the spin-restricted, spin-unrestricted, and spin-restricted open-shell variants of HF and Kohn–Sham density functional theory (DFT). TeraChem works with atom-centered Gaussian basis sets through  $d$  angular momenta and offers effective core potentials to account for core electrons in heavy atoms.<sup>39,40</sup> It can use either Cartesian or spherical basis representations (the latter by projection) for the atomic orbitals, and both conventional and incremental Fock matrix<sup>41</sup> constructions are available for HF and DFT.

TeraChem encompasses a variety of common exchange–correlation functionals, including local spin density approximation functionals, generalized gradient approximation functionals, hybrid functionals, and range-separated functionals, and both static and dynamic (i.e., multigrid) grids are available for DFT. DFT+U<sup>42,43</sup> as well as Grimme's D2<sup>44</sup> and D3<sup>45,46</sup> dispersion and geometric counterpoise<sup>47</sup> corrections are implemented. All integral contractions are formulated in a generalized global plus range-separated form; thus, TeraChem also provides solutions to the Coulomb-attenuated Schrödinger equation.<sup>48,49</sup> Density functional tight binding<sup>50–52</sup> (DFTB) energies and gradients with the 3ob-3-1 Slater–Koster parameter set<sup>53–56</sup> are available through an interface to the DFTB+ package.<sup>57</sup> The GFN-xTB<sup>58</sup> and GFN2-xTB<sup>59</sup> semiempirical tight-binding methods are available internally through a new semiempirical integral library, which provides semiempirical equivalents to the aforementioned core ab initio electronic structure quantities.<sup>60</sup>

Robust behavior of the SCF procedure is often dependent on the quality of the starting wavefunction and convergence acceleration algorithms. TeraChem provides the projection from a minimal basis calculation, the superposition of atomic densities method,<sup>41,61</sup> the maximum orbital overlap method,<sup>62</sup> and several fragment-based schemes for generating trial wave functions. Pulay's direct inversion of the iterative subspace<sup>63</sup> (DIIS) and a hybrid DIIS+ADIIS scheme<sup>64</sup> are implemented as convergence accelerators. TeraChem also includes level-shifting<sup>65,66</sup> and fractional occupation number<sup>67</sup> (FON) strategies to assist with convergence in difficult cases, and a FON annealing method is provided to allow exploration in the space of possible electronic solutions. Orbital-free density matrix-based SCF calculations, that is, without Fock matrix diagonalization, based on a GPU-accelerated implementation of Niklasson's canonical fourth-order trace-resetting purification approach are available as well.<sup>68</sup>

TeraChem can also compute a host of electronic and vibrational properties, including dipole vectors, polarizability tensors, electrostatic potentials (ESPs), electronic density maps, and bond order matrices. In addition to Mulliken and

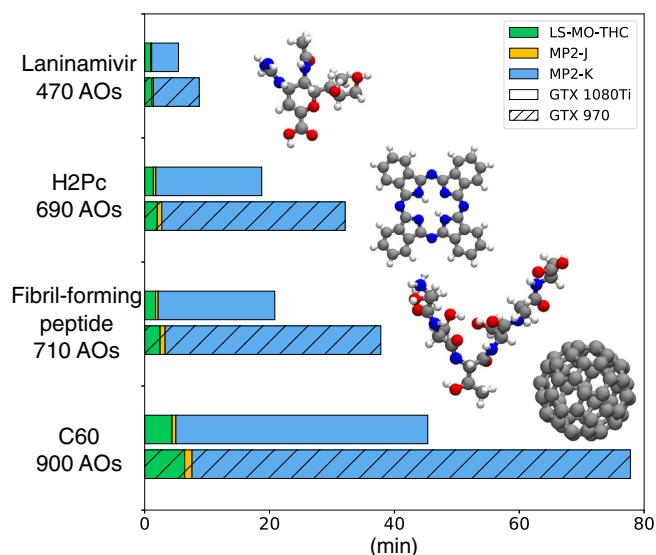
Voronoi deformation density<sup>69</sup> population analysis, TeraChem is interfaced with NBO 6.0<sup>70</sup> for more advanced natural bond order (NBO) and natural population analyses. Numerical Hessians can be used to perform thermochemical analysis or sample initial geometries from either the Husimi or Wigner distributions (in the harmonic approximation at zero or finite temperature) for dynamics simulations and are available for all methods with analytical gradients. Electronic absorption spectra can be calculated with single- or multi-reference methods such as TDDFT and CASPT2. Additionally, TeraChem Molden outputs from spherical basis SCF calculations can be used to run simplified TD-DFT<sup>71,72</sup> calculations with the sTDA program<sup>73</sup> to obtain UV-Vis absorption and electronic circular dichroism spectra.

### 3 | SINGLE REFERENCE POST-SCF METHODS

Excited state energies, gradients, and response properties can be calculated using TDHF, TDDFT, or their TDA variants, CIS and TDA-TDDFT.<sup>16</sup> Spin-flip<sup>74–76</sup> variants of the latter two single excitation methods are also available, although these methods are known to suffer from spin contamination issues.<sup>77</sup> In order to overcome these issues, the hole–hole TDA (hh-TDA) method, similar in spirit to the recent particle–particle TDA method,<sup>78–80</sup> has recently been implemented and shows promising results for nonadiabatic dynamics simulations.<sup>81,82</sup> TeraChem provides energies and analytic gradients for THC-MP2<sup>31</sup> and THC-SOS-MP2<sup>31–33</sup> through a THC module, with demonstrative timings shown in Figure 3 for least-squares THC fit to ERIs in the molecular orbital basis (LS-THC-MO). Unlike density fitting<sup>83–85</sup> (which lowers the computational time but does not affect the scaling behavior for MP2), THC formally reduces the scaling of MP2 from  $O(N^5)$  to  $O(N^4)$ .<sup>30</sup> The SOS-MP2 method<sup>86</sup> neglects exchange-like terms to arrive at a formal scaling of  $O(N^4)$ , and THC-SOS-MP2 further reduces this scaling to  $O(N^3)$ . Recently, a GPU-accelerated coupled-cluster code has also been implemented in TeraChem, enabling coupled-cluster singles and doubles (CCSD) as well as any method that can be written as a subset of CCSD diagrams.<sup>87</sup>

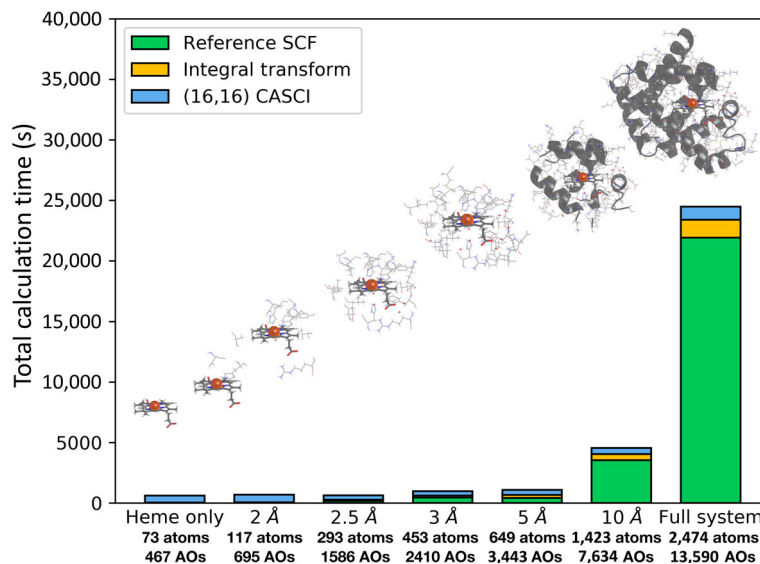
### 4 | MULTI-REFERENCE METHODS

TeraChem's direct determinantal CI library enables CASCI and CASSCF calculations with upwards of  $10^9$  determinants on a single GPU, as showcased in Figure 4. Floating occupation molecular orbital CASCI<sup>20</sup> (FOMO-CASCI) and CIS natural orbital CASCI<sup>88</sup> (CISNO-CASCI) are available as low-cost, robust alternatives to SA-CASSCF. TeraChem's state-of-the-art atomic orbital basis SA-CASSCF<sup>19,24,26</sup> and  $\alpha$ -CASSCF<sup>89</sup> implementations offer energies, gradients, and nonadiabatic couplings with effective quadratic scaling with respect to the number of orbitals (assuming an active space of constant size); as with the SCF methods, this scaling is due to the element sparsity present in the AO basis. Spin–orbit couplings are available for the CASCI and CASSCF methods. Recently implemented electronic structure methods include reduced rank full CI<sup>90</sup> and time-dependent CASCI,<sup>91</sup> while recent algorithmic advances include the use of



**FIGURE 3** Computational wall times of THC-MP2 calculations, where bar colors represent different components, that is, constructing the LS-THC-MO tensors, MP2 Coulomb-like energies, and MP2 exchange-like energies. Calculations use cc-pVDZ basis set, cc-pVDZ-RI auxiliary basis set for density fitting, and THC grids optimized for cc-pVDZ. Timings were run using a single GTX 980 or 1080Ti GPU and a single thread of a 3.33 GHz Intel Xeon X5680 CPU

**FIGURE 4** Timings (in seconds) for hybrid QM/MM calculations of myoglobin (PDB ID: 3RGK) at the HF-CAS-(16,16)-CI/6-31G level of theory using a single V100 GPU and a 3.4 GHz Intel Xeon E5-2643v4 CPU. The various QM regions were carved out by including entire residues within a given distance of the heme cofactor. Note that the CASCI portion of the calculation dominates for small QM sizes but remains fairly constant and is heavily outweighed by the SCF procedure for the entire protein



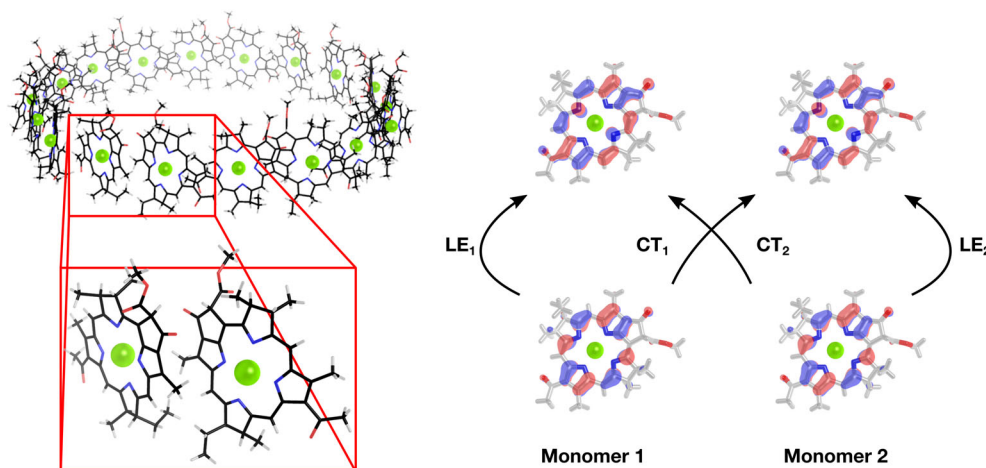
multiple GPUs for CI,<sup>92</sup> mixed precision during  $\sigma$  builds,<sup>93</sup> and an efficient configuration-state function to determinant transform.<sup>94</sup> TeraChem also offers density matrix renormalization group<sup>95</sup> calculations through an interface developed by Keller to a GPU-accelerated version of QCMarquis.<sup>96</sup>

Several methods that treat both static and dynamic correlation efficiently are incorporated within TeraChem. The embedding-based CASCI/DFT hybrid approach developed by Pijeu and Hohenstein is implemented for FOMO-CASCI, CISNO-CASCI and SA-CASSCF wave functions.<sup>97</sup> Spin-restricted ensemble-reference Kohn-Sham (REKS) energies and analytic gradients are available for 2 electrons in 2 orbitals (REKS(2,2)), state-averaged REKS(2,2) (SA-REKS(2,2)), and state-interaction SA-REKS(2,2) (SI-SA-REKS(2,2)), with nonadiabatic couplings also implemented for SI-SA-REKS(2,2).<sup>98,99</sup> Energies are available for the 4 electrons in 4 orbitals variants, with gradients only being implemented for REKS(4,4) and SA-REKS(4,4).<sup>100</sup> State-specific THC-CASPT2 is implemented using the supporting subspace technique to leverage the existing THC-MP2 routines, ensuring that THC-CASPT2 energies only scale as  $O(N^4)$  with respect to the number of molecular orbitals (for fixed active space size).<sup>36</sup>

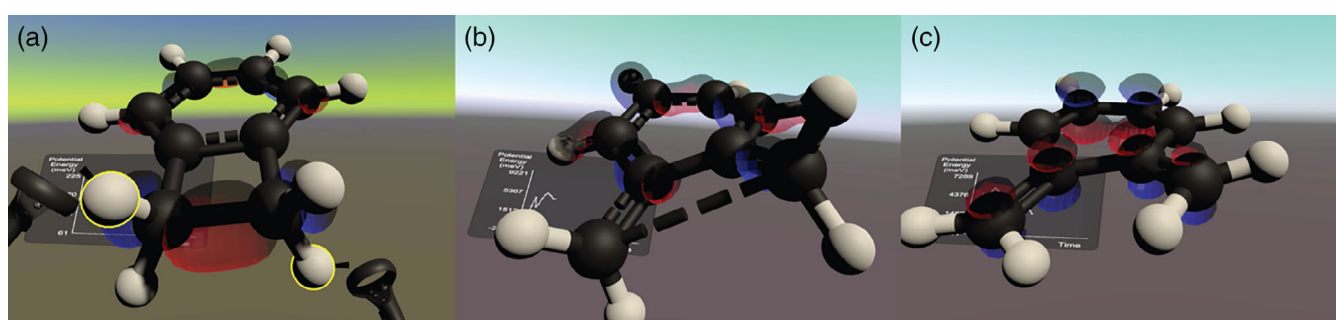
## 5 | ENVIRONMENT EFFECTS

Although TeraChem can tackle large scale electronic structure calculations on desktop hardware, many systems of interest, such as protein complexes, are simply too large for a purely ab initio treatment; as a result, TeraChem includes a variety of methods to include environmental effects. Energies and gradients for implicit solvation<sup>101</sup> are provided through a GPU-accelerated conductor-like polarizable continuum model (C-PCM) implementation with the improved-switching Gaussian approach.<sup>102</sup> The development of C-PCM gave rise to two new core electronic structure operations for one-electron integrals: the construction of  $\mathbf{c}$ , which describes the solvent cavity surface charge-solute density interaction, and  $\Delta\mathbf{F}^S$ , the solvent contribution to the Fock matrix.<sup>37</sup> C-PCM is available in conjunction with HF and DFT in the ground state, and state-specific and linear-response versions are available with TDA-TDDFT. Encapsulation of  $\mathbf{c}$  and  $\Delta\mathbf{F}^S$  as building blocks, which may find additional use in constructing ESPs, restricted ESP atomic charge fitting, or embedding schemes, is ongoing.<sup>103</sup>

Explicit MM waters can be modeled using internal implementations of the SPC, TIP3P, TIP4P, SWM4, SWM4-DP, or SWM4-NDP force fields.<sup>104,105</sup> More general QM/MM simulations can be carried out through a file-based AMBER interface<sup>106</sup> or with a statically linked OpenMM 7.0.<sup>107</sup> Fragment-based approaches are an attractive alternative to QM/MM; as a result, TeraChem also includes energies and gradients for an ab initio exciton model with both point-dipole and full dimer couplings between monomers. The exciton model has been successfully applied to multichromophoric structures such as light-harvesting system II (LH2), as depicted in Figure 5, enabling excited state ab initio molecular dynamics (AIMD) with more than 3,000 atoms.<sup>108–110</sup>



**FIGURE 5** The ab initio exciton model with locally excited (LE) and charge transfer (CT) states applied to the 18 BChl-*a* chromophore B850 assembly of the light harvesting system II (LH2). In recent work, Li et al.<sup>110</sup> benchmarked against TDDFT and EOM-CC2 calculations of the supersystem to show the robustness and accuracy of the exciton model



**FIGURE 6** Mechanochemical ring opening of benzocyclobutene using real-time interactive molecular dynamics with GFN2-xTB. Dynamic bond orders and molecular orbitals show (a) the sigma bond in the reactant, (b) the bond rearrangement at the transition state, and (c) the new  $\pi$  orbital in the product

## 6 | GEOMETRY OPTIMIZATIONS AND MOLECULAR DYNAMICS

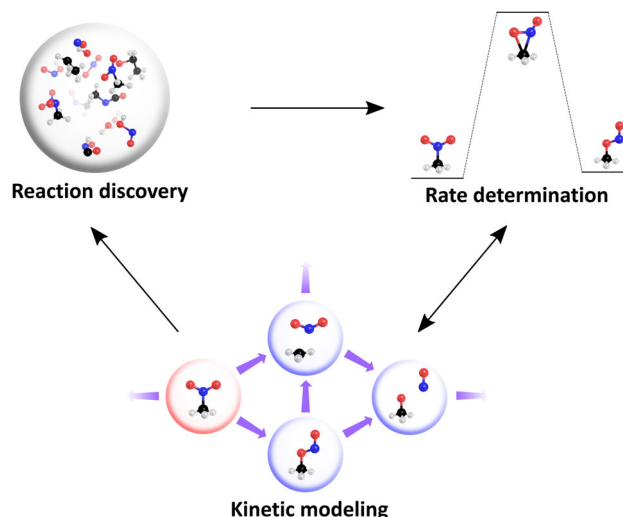
TeraChem development follows the philosophy of separating electronic structure and driver (e.g., optimization, molecular dynamics) codes; however, several driver-level codes are included and available to users through the input file. TeraChem is interfaced with DL-FIND<sup>111,112</sup> to enable geometry minimizations and transition state searches and with geomeTRIC<sup>113</sup> to enable geometry minimizations with translational-rotational invariant coordinates; therefore, TeraChem can perform standalone minimum energy pathway optimizations using nudged elastic band (NEB), climbing image NEB (CI-NEB), and the simplified string method.<sup>114–117</sup> Constraints on atomic position, bond length, angle, torsion, and molecular translation/rotation degrees of freedom can be applied during energy minimizations.

Nonperiodic AIMD is available for NVE and NVT ensembles, and includes velocity rescaling,<sup>118</sup> Nose-Hoover chains,<sup>119</sup> and Langevin<sup>120</sup> thermostats. Niklasson's density propagation schemes are used to ensure time reversibility of the electronic degrees of freedom,<sup>121,122</sup> and a two-level multiple timestep scheme<sup>123</sup> is supported with DFTB inner timesteps. Time-dependent spherical, hemispherical, disk, and surface boundary conditions can be applied to all AIMD simulations. TeraChem is interfaced with PLUMED<sup>124</sup> to enable enhanced sampling techniques, such as metadynamics and umbrella sampling, and has implementations for boxed molecular dynamics,<sup>125</sup> adaptive hyperdynamics,<sup>126</sup> and steered molecular dynamics.<sup>127</sup> Real-time interactive AIMD can be performed on systems with a few dozen atoms, as demonstrated in Figure 6.<sup>128</sup>

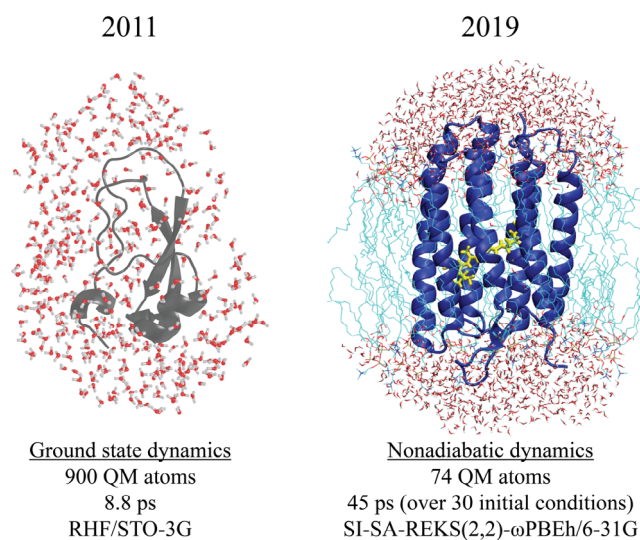
Geometry optimizations and AIMD simulations enable several downstream quantum chemistry workloads. One such workload is automated reaction discovery, where accelerated AIMD is first used to discover chemical reactions and then geometry and MEP optimizations are used to refine chemical species and estimate barrier heights.<sup>129–131</sup> For reaction discovery, artificial forces are injected into a high temperature AIMD simulation to encourage chemical reactions; often, these simulations use robust SCF convergence options such as the aforementioned FON annealing to ensure the stability of the trajectories. During refinement, discovered species are optimized separately and minimum



**FIGURE 7** Schematic overview of the ab initio nanoreactor framework. Reaction networks are iteratively built up through three phases: reaction discovery through accelerated molecular dynamics, rate determination through minimum energy pathway optimization, and running kinetic models to generate new concentration profiles for discovery runs and pinpoint rate-limiting intermediates for further refinement



**FIGURE 8** TeraChem's capability to simulate ab initio molecular dynamics for proteins as GPU hardware and algorithms improve, from ground state dynamics and optimization in 2011<sup>135,136</sup> to multireference nonadiabatic dynamics in 2019<sup>141</sup>



energy pathways are computed for each reaction. Combining thermodynamic information obtained during refinement with transition state theory provides estimated rates for the construction of a kinetic model, enabling the simulation of system evolution on a significantly longer timescale. Both the discovery and refinement stages of this workflow, shown schematically in Figure 7, benefit heavily from GPU-accelerated electronic structure. As a result, one can use ab initio calculations to glean insights into complex reaction networks such as those involved in the origins of life.<sup>129,132–134</sup>

The development of TeraChem enabled ground state dynamics<sup>135</sup> and geometry optimization<sup>136</sup> on entire proteins at the HF and DFT levels of theory. These calculations made it possible for the development of methods that relied on ab initio protein structures, such as crystallographic refinement<sup>137,138</sup> and ligand binding affinities.<sup>139</sup> Using TeraChem, it is becoming routine to run QM/MM nonadiabatic dynamics simulations on proteins.<sup>140,141</sup> Typically, the QM region contains tens to hundreds of atoms, and several dozen trajectories can be run for hundreds of femtoseconds using a multireference correlation method, as showcased in Figure 8.

An interface to the FMS90 code enables TeraChem to drive large-scale ab initio multiple spawning (AIMS) simulations for studying photochemical processes with nonadiabatic dynamics.<sup>142</sup> Unlike geometry optimization and AIMD, nonadiabatic dynamics requires time-dependent electronic structure information to be exchanged between TeraChem and FMS90. Wavefunction intermediates (e.g., molecular orbitals, CI vectors) are passed back to TeraChem after each timestep to ensure continuity of electronic wavefunctions (especially consistency of the phase) over the course of the trajectory. AIMS simulations are enabled for the CIS/TDA-TDDFT, hh-TDA, FOMO-CASCI, CISNO-CASCI, SA-CAS-SCF,  $\alpha$ -SA-CASSCF, and SI-SA-REKS(2,2) methods, and have been used to study systems such as provitamin D3,<sup>25</sup> 4-(*N,N*-dimethylamino)benzotrile,<sup>143</sup> ethylene, methaniminium cation, malonaldehyde,<sup>144</sup> ultrafast electron diffraction

Method	Performance (ps/day)
HF	26.02
B3LYP	12.74
CIS	4.41
TDA-TDDFT/ $\omega(0.3)$ -PBEh	2.39
hh-TDA/ $\omega(0.3)$ -PBEh	2.02
FOMO-CASCI(2,2)	4.56
SA2-CASSCF(2,2)	3.33
SI-SA-REKS(2,2)/ $\omega(0.2)$ -PBEh	0.22

**TABLE 1** Performance estimates for ab initio molecular dynamics based on 50 fs runs

Note: Simulations used a 0.5 fs timestep with the 6-31G\* basis set for *cis*-stilbene (26 atoms/234 basis functions) on one NVIDIA Tesla V100 GPU and a single core of an Intel Xeon 3.4 GHz E5-2643v4 CPU with TeraChem. All excited state methods calculate two roots and use FMS90/TeraChem to run a single trajectory basis function (TBF) on the first excited state with no adaptive timestepping. AIMS simulations generally decrease the timestep in regions of nonadiabatic coupling and usually involve multiple TBFs when nonadiabatic effects are modeled. These factors can decrease the simulation throughput.

experiments in cyclohexadiene<sup>145</sup> and 1,2-diiodotetrafluoroethane,<sup>146</sup> the retinal protonated Schiff base in channelrhodopsin 2<sup>140</sup> and bacteriorhodopsin,<sup>141</sup> and *cis*-stilbene.<sup>147</sup> Table 1 summarizes performance estimates for both ground state and nonadiabatic molecular dynamics with the variety of electronic structure methods available in TeraChem on modern scientific-grade GPUs.

## 7 | CONCLUSION

A decade ago, TeraChem was the first quantum chemistry code to leverage GPUs as an alternative computing architecture for electronic structure. The construction of Coulomb and exchange matrices can be done extremely efficiently on the GPU, especially when combined with AO basis sparsity and mixed precision techniques. As functionality was added to TeraChem, efficient electronic structure calculations were enabled by reusing these electronic structure subroutines; however, it has since become clear that these integral contractions were the first of several core building blocks for electronic structure, which now also include generalized OPDMs/TPDMs, sigma builds, and two cavity-density interactions. Just as BLAS operations revolutionized numerical linear algebra, electronic structure methods can leverage GPUs efficiently by feeding generalized density matrices into these highly optimized building blocks.

From a development standpoint, the rise of core electronic structure operations also provides a clear target for further improvement. Code generation techniques can be used to provide support for higher angular momentum functions<sup>148,149</sup> and will open up a whole new class of chemical systems for study with TeraChem. Hardening the interfaces will allow more complex compositions of primitives and lead to new electronic structure methods. Lastly, and perhaps most excitingly, there could still be further performance gains to be made by exploiting new hardware accelerators. The new focus for GPU hardware development has shifted towards tensor cores and half precision support for machine learning applications and real-time ray tracing for graphics rendering pipelines. Traditional GPU programming paradigms are also evolving. NVIDIA's Volta generation enables greater control over individual threads in a warp and high bandwidth solutions like NVLink and NVSwitch help mitigate expensive CPU-GPU or direct GPU-GPU memory transfers. Utilizing these new features for electronic structure software remains an active area of research within the community. While GPUs were a clear contender in the previous decade due to the close connection to linear algebra and rise of computer graphics, the next decade is full of unique architectures that may be useful to computational chemistry. For example, one can imagine using tensor processing units to compute a stack of matrix builds or using field-programmable gate arrays as a stepping stone to develop an application-specific integrated circuit for electronic structure. The modularization strategy of identifying, encapsulating, and reusing core operations ensures that all electronic structure methods benefit immediately from any algorithmic development on these novel computing architectures.

## ACKNOWLEDGMENTS

The authors would like to thank Leah I. Bendavid, Andrew Durden, Ethan Curtis, Lin Fan, Jason E. Ford, William J. Glover, Jeffrey R. Gour, Ruben Guerrero, Daniel Hollas, Tim Ioannidis, K. Grace Johnson, Sebastian Keller, Heather J. Kulik, Dean Lahana, Ruibin Liang, Brendan D. Mar, Michael S. Miller, Michael O'Connor, Robert M. Parrish, Wei-Tao Peng, Helena W. Qi, Torsten Sachse, Aaron Sisto, Yuanheng Wang, Hayley Weir, Jimmy K. Yu, and Xiaolei Zhu for significant contributions to the TeraChem codebase, and Alice Walker for discussions on the QM/MM setup for myoglobin. This work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program and ONR N000-14-17-12875. S. S. acknowledges support from an NSF Graduate Research Fellowship and Boeing through the Stanford SystemX Alliance. C. B. was supported by the German National Academy of Sciences Leopoldina through the Leopoldina Fellowship Program (Project No. LPDS 2018-09).

## CONFLICT OF INTEREST

T. J. M. and I. S. U. are cofounders of PetaChem, LLC.

## AUTHOR CONTRIBUTIONS

**Stefan Seritan:** Writing-original draft; writing-review and editing. **Christoph Bannwarth:** Writing-original draft; writing-review and editing. **Bryan Fales:** Writing-original draft; writing-review and editing. **Edward Hohenstein:** Writing-original draft; writing-review and editing. **Christine Isborn:** writing-review and editing. **Sara Kokkila-Schumacher:** writing-review and editing. **Xin Li:** writing-review and editing. **Fang Liu:** writing-review and editing. **Nathan Luehr:** writing-review and editing. **James Snyder:** writing-review and editing. **Chenchen Song:** Writing-original draft; writing-review and editing. **Alexey Titov:** writing-review and editing. **Ivan Ufimtsev:** writing-review and editing. **Lee-Ping Wang:** writing-review and editing. **Todd Martinez:** Conceptualization; funding acquisition; writing-original draft; writing-review and editing.

## ORCID

Stefan Seritan  <https://orcid.org/0000-0002-8808-6886>  
Christoph Bannwarth  <https://orcid.org/0000-0003-3242-496X>  
Bryan S. Fales  <https://orcid.org/0000-0002-0627-2782>  
Edward G. Hohenstein  <https://orcid.org/0000-0002-2119-2959>  
Sara I. L. Kokkila-Schumacher  <https://orcid.org/0000-0002-2338-4815>  
Xin Li  <https://orcid.org/0000-0001-6508-8355>  
Fang Liu  <https://orcid.org/0000-0003-1322-4997>  
Nathan Luehr  <https://orcid.org/0000-0001-8740-6211>  
Chenchen Song  <https://orcid.org/0000-0001-6294-2253>  
Ivan S. Ufimtsev  <https://orcid.org/0000-0003-2632-0988>  
Lee-Ping Wang  <https://orcid.org/0000-0003-3072-9946>  
Todd J. Martinez  <https://orcid.org/0000-0002-4798-8947>

## RELATED WIRES ARTICLES

[MOLCAS-a software for multiconfigurational quantum chemistry calculations](#)

[Q-Chem: an engine for innovation](#)

[Turbomole](#)

[BAGEL: Brilliantly Advanced General Electronic-Structure Library](#)

[Molpro: a general purpose quantum chemistry program package](#)

## REFERENCES

1. Levine B, Martinez TJ. Hijacking the playstation2 for computational chemistry. *Abs Papers ACS*. 2003;226:U426.
2. NVIDIA Corporation. CUDA C Programming Guide. <https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html> (accessed August 24, 2018).
3. Ufimtsev IS, Martinez TJ. Graphical processing units for quantum chemistry. *Comput Sci Eng*. 2008;10:26–34.
4. Ufimtsev IS, Martinez TJ. Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J Chem Theory Comput*. 2008;4:222–231.

5. Ufimtsev IS, Martínez TJ. Quantum chemistry on graphical processing units. 2. Direct self-consistent-field implementation. *J Chem Theory Comput.* 2009;5:1004–1015.
6. Ufimtsev IS, Martínez TJ. Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J Chem Theory Comput.* 2009;5:2619–2628.
7. Luehr N, Ufimtsev IS, Martínez TJ. Dynamic precision for electron repulsion integral evaluation on graphical processing units (GPUs). *J Chem Theory Comput.* 2011;7:949–954.
8. Stone JE, Hardy DJ, Ufimtsev IS, Schulten K. GPU-accelerated molecular modeling coming of age. *J Mol Graph Model.* 2010;29:116–125.
9. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *WIREs Comput Mol Sci.* 2013;3:198–210.
10. Zheng M, Li X, Guo L. Algorithms of GPU-enabled reactive force field (ReaxFF) molecular dynamics. *J Mol Graph Model.* 2013;41:1–11.
11. Kutzner C, Páll S, Fechner M, Esztermann A, de Groot BL, Grubmüller H. Best bang for your buck: GPU nodes for GROMACS biomolecular simulations. *J Comput Chem.* 2015;36:1990–2008.
12. Evan JA, Charles LB. Efficient implementation of constant pH molecular dynamics on modern graphics processors. *J Comput Chem.* 2016;37:2171–2180.
13. Almlöf J, Faegri K, Korsell K. Principles of a direct SCF approach to LCAO-MO ab-initio calculations. *J Comput Chem.* 1982;3:385–399.
14. Tomov S, Dongarra J, Baboulin M. Towards dense linear algebra for hybrid GPU accelerated manycore systems. *Parallel Comput.* 2010;36:232–240.
15. Titov AV, Ufimtsev IS, Luehr N, Martínez TJ. Generating efficient quantum chemistry codes for novel architectures. *J Chem Theory Comput.* 2013;9:213–221.
16. Isborn CM, Luehr N, Ufimtsev IS, Martínez TJ. Excited-state electronic structure with configuration interaction singles and Tamm-Dancoff time-dependent density functional theory on graphical processing units. *J Chem Theory Comput.* 2011;7:1814–1823.
17. Foresman JB, Head-Gordon M, Pople JA, Firisch MJ. Toward a systematic molecular orbital theory for excited states. *J Phys Chem.* 1992;96:135–149.
18. Casida ME. Time-dependent density functional response theory for molecules. In: Chong DP, editor. *Recent advances in density functional methods.* Singapore: World Scientific, 1995; p. 155–192.
19. Hohenstein EG, Luehr N, Ufimtsev IS, Martínez TJ. An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units. *J Chem Phys.* 2015;142:224103.
20. Hohenstein EG, Bouduban MEF, Song C, Luehr N, Ufimtsev IS, Martínez TJ. Analytic first derivatives of floating occupation molecular orbital-complete active space configuration interaction on graphical processing units. *J Chem Phys.* 2015;143:014111.
21. Hohenstein EG. Analytic formulation of derivative coupling vectors for complete active space configuration interaction wavefunctions with floating occupation molecular orbitals. *J Chem Phys.* 2016;145:174110.
22. Hohenstein EG. Mechanism for the enhanced excited-state Lewis acidity of methyl viologen. *J Am Chem Soc.* 2016;138:1868–1876.
23. Fales BS, Levine BG. Nanoscale multireference quantum chemistry: Full configuration interaction on graphical processing units. *J Chem Theory Comput.* 2015;11:4708–4716.
24. Snyder JW, Hohenstein EG, Luehr N, Martínez TJ. An atomic orbital-based formulation of analytical gradients and nonadiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units. *J Chem Phys.* 2015;143:154107.
25. Snyder JW, Curchod BFE, Martínez TJ. GPU-accelerated state-averaged complete active space self-consistent field interfaced with ab initio multiple spawning unravels the photodynamics of provitamin D3. *J Phys Chem Lett.* 2016;7:2444–2449.
26. Snyder JW, Fales BS, Hohenstein EG, Levine BG, Martínez TJ. A direct-compatible formulation of the coupled perturbed complete active space self-consistent field equations on graphical processing units. *J Chem Phys.* 2017;146:174113.
27. Hohenstein EG, Parrish RM, Martínez TJ. Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory. *J Chem Phys.* 2012;137:044103.
28. Parrish RM, Hohenstein EG, Martínez TJ, Sherrill CD. Tensor hypercontraction. II. Least-squares renormalization. *J Chem Phys.* 2012;137:224106.
29. Hohenstein EG, Parrish RM, Sherrill CD, Martínez TJ. Communication: Tensor hypercontraction. III. Least-squares tensor hypercontraction for the determination of correlated wavefunctions. *J Chem Phys.* 2012;137:221101.
30. Kokkila Schumacher SIL, Hohenstein EG, Parrish RM, Wang L-P, Martínez TJ. Tensor hypercontraction second-order Møller-Plesset perturbation theory: Grid optimization and reaction energies. *J Chem Theory Comput.* 2015;11:3042–3052.
31. Song C, Martínez TJ. Analytical gradients for tensor hyper-contracted MP2 and SOS-MP2 on graphical processing units. *J Chem Phys.* 2017;147:161723.
32. Song C, Martínez TJ. Atomic orbital-based SOS-MP2 with tensor hypercontraction. I. GPU-based tensor construction and exploiting sparsity. *J Chem Phys.* 2016;144:174111.
33. Song C, Martínez TJ. Atomic orbital-based SOS-MP2 with tensor hypercontraction. II. Local tensor hypercontraction. *J. Chem. Phys.* 2017;146:034104.
34. Hohenstein EG, Kokkila SIL, Parrish RM, Martínez TJ. Quartic scaling second-order approximate coupled cluster singles and doubles via tensor hypercontraction: THC-CC2. *J Chem Phys.* 2013;138:124111.

35. Hohenstein EG, Kokkila SIL, Parrish RM, Martínez TJ. Tensor hypercontraction equation-of-motion second-order approximate coupled cluster: Electronic excitation energies in  $O(N^4)$  time. *J Phys Chem B*. 2013;117:12972–12978.
36. Song C, Martínez TJ. Reduced scaling CASPT2 using supporting subspaces and tensor hyper-contraction. *J Chem Phys*. 2018;149:044108.
37. Liu F, Luehr N, Kulik HJ, Martínez TJ. Quantum chemistry for solvated molecules on graphical processing units using polarizable continuum models. *J Chem Theory Comput*. 2015;11:3131–3144.
38. Seritan S, Thompson K, Martínez TJ. TeraChem cloud: A high-performance computing service for scalable distributed GPU-accelerated electronic structure calculations. *J Chem Inf Model*. 2020;60:2126–2137.
39. Song C, Wang LP, Sachse T, Preiß J, Presselt M, Martínez TJ. Efficient implementation of effective core potential integrals and gradients on graphical processing units. *J Chem Phys*. 2015;143:014114.
40. Song C, Wang L-P, Martínez TJ. Automated code engine for graphical processing units: Application to the effective Core potential integrals and gradients. *J Chem Theory Comput*. 2016;12:92–106.
41. Almlöf J, Faegri K, Korsell K. Principles for a direct SCF approach to LCAO-MO ab-initio calculations. *J Comput Chem*. 1982;3:385–399.
42. Anisimov VI, Zaanen J, Andersen OK. Band theory and Mott insulators: Hubbard U instead of Stoner I. *Phys Rev B*. 1991;44:943–954.
43. Cococcioni M, de Gironcoli S. Linear response approach to the calculation of the effective interaction parameters in the LDA + U method. *Phys Rev B*. 2005;71:035105.
44. Grimme S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem*. 2006;27:1787–1799.
45. Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys*. 2010;132:154104.
46. Grimme S, Ehrlich S, Goerigk L. Effect of the damping function in dispersion corrected density functional theory. *J Comput Chem*. 2011;32:1456–1465.
47. Kruse H, Grimme S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree–Fock and density functional theory calculations for large systems. *J Chem Phys*. 2012;136:154101.
48. Adamson RD, Dombroski JP, Gill PMW. Chemistry without Coulomb tails. *Chem Phys Lett*. 1996;254:329–336.
49. Nolen JA Jr, Schiffer JP. Coulomb energies. *Annu Rev Nucl Sci*. 1969;19:471–526.
50. Elstner M, Porezag D, Jungnickel G, et al. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B*. 1998;58:7260–7268.
51. Porezag D, Frauenheim T, Kohler T, Seifert G, Kaschner R. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys Rev B*. 1995;51:12947–12957.
52. Seifert G, Porezag D, Frauenheim T. Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme. *Int J Quantum Chem*. 1996;58:185–192.
53. Gaus M, Goez A, Elstner M. Parametrization and benchmark of DFTB3 for organic molecules. *J Chem Theory Comput*. 2013;9:338–354.
54. Gaus M, Lu X, Elstner M, Cui Q. Parameterization of DFTB3/3OB for sulfur and phosphorus for chemical and biological applications. *J Chem Theory Comput*. 2014;10:1518–1537.
55. Kubillus M, Kubař T, Gaus M, Řezáč J, Elstner M. Parameterization of the DFTB3 method for Br, Ca, Cl, F, I, K, and Na in organic and biological systems. *J Chem Theory Comput*. 2015;11:332–342.
56. Lu X, Gaus M, Elstner M, Cui Q. Parameterization of DFTB3/3OB for magnesium and zinc for chemical and biological applications. *J Phys Chem B*. 2015;119:1062–1082.
57. Aradi B, Hourahine B, Frauenheim T. DFTB+, a sparse matrix-based implementation of the DFTB method. *J Phys Chem A*. 2007;111:5678–5684.
58. Grimme S, Bannwarth C, Shushkov P. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements ( $Z = 1–86$ ). *J Chem Theory Comput*. 2017;13:1989–2009.
59. Bannwarth C, Ehlert S, Grimme S. GFN2-xTB – An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J Chem Theory Comput*. 2019;15:1652–1671.
60. Bannwarth C, Martínez TJ. Novel method combinations enabled by interfacing a semiempirical integral library to a modular ab initio electronic structure framework. 2020 (in preparation).
61. Van Lenthe JH, Zwaans R, Van Dam HJJ, Guest MF. Starting SCF calculations by superposition of atomic densities. *J Comput Chem*. 2006;27:926–932.
62. Langlois J-M, Yamasaki T, Muller RP, Goddard WA. Rule-based trial wave functions for generalized valence bond theory. *J Phys Chem*. 1994;98:13498–13505.
63. Pulay P. Improved SCF convergence acceleration. *J Comput Chem*. 1982;3:556–560.
64. Hu X, Yang W. Accelerating self-consistent field convergence with the augmented Roothaan–Hall energy function. *J Chem Phys*. 2010;132:054109.
65. Saunders VR, Hillier IH. A “level-shifting” method for converging closed shell Hartree–Fock wave functions. *Int J Quantum Chem*. 1973;7:699–705.

66. Cancès E, Le Bris C. On the convergence of SCF algorithms for the Hartree–Fock equations. *ESAIM Math Model Numer.* 2000;34:749–774.
67. Rabuck AD, Scuseria GE. Improving self-consistent field convergence by varying occupation numbers. *J Chem Phys.* 1998;110:695.
68. Niklasson AMN, Tymczak CJ, Challacombe M. Trace resetting density matrix purification in  $O(N)$  self-consistent-field theory. *J Chem Phys.* 2003;118:8611–8620.
69. Fonseca Guerra C, Handgraaf J-W, Jan Baerends E, Matthias Bickelhaupt F. Voronoi deformation density (VDD) charges: Assessment of the Mulliken methods for charge analysis. *J Comput Chem.* 2004;25:189–210.
70. Glendening ED, Badenhop JK, Reed AE, et al. NBO 6.0, Theoretical Chemistry Institute, University of Wisconsin, Madison; 2013.
71. Grimme S. A simplified Tamm–Dancoff density functional approach for the electronic excitation spectra of very large molecules. *J Chem Phys.* 2013;138:244104.
72. Bannwarth C, Grimme S. A simplified time-dependent density functional theory approach for electronic ultraviolet and circular dichroism spectra of very large molecules. *Comput Theoret Chem.* 2014;1040-1041:45–53.
73. de Wergifosse M, Bannwarth C, Ehler S, Grimme S. grimme-lab/stda: stda program for computing excited states and response functions via simplified TD-DFT methods (sTDA, sTD-DFT, SF-STD-DFT). <https://github.com/grimme-lab/stda> (accessed December 24, 2019).
74. Krylov AI. Size-consistent wave functions for bond-breaking: The equation-of-motion spin-flip model. *Chem Phys Lett.* 2001;338:375–384.
75. Krylov AI. Spin-flip configuration interaction: An electronic structure model that is both variational and size-consistent. *Chem Phys Lett.* 2001;350:522–530.
76. Shao Y, Head-Gordon M, Krylov AI. The spin-flip approach within time-dependent density functional theory: Theory and applications to diradicals. *J Chem Phys.* 2003;118:4907–4818.
77. Sears JS, Sherrill CD, Krylov AI. A spin-complete version of the spin-flip approach to bond breaking: What is the impact of obtaining spin eigenfunctions? *J Chem Phys.* 2003;118:9084–9094.
78. Yang Y, Van Aggelen H, Yang W. Double, Rydberg and charge transfer excitations from pairing matrix fluctuation and particle–particle random phase approximation. *J Chem Phys.* 2013;139:224105.
79. Van Aggelen H, Yang Y, Yang W. Exchange-correlation energy from pairing matrix fluctuation and the particle-particle random-phase approximation. *Phys Rev A.* 2013;88:030501(R).
80. Peng D, Van Aggelen H, Yang Y, Yang W. Linear-response time-dependent density-functional theory with pairing fields. *J Chem Phys.* 2014;140:18A522.
81. Bannwarth C, Yu JK, Hohenstein EG, Martínez TJ. Hole–hole Tamm–Dancoff-approximated density functional theory: A highly efficient electronic structure method incorporating dynamic and static correlation. *ChemRxiv.* 2020;chemrxiv.11828256.v2: 1–35.
82. Yu JK, Bannwarth C, Hohenstein EG, Martínez TJ. Ab initio multiple spawning dynamics with the hole–hole Tamm–Dancoff approximation. 2020 (in preparation).
83. Vahtras O, Almlöf J, Feyereisen M. Integral approximations for LCAO-SCF calculations. *Chem Phys Lett.* 1993;213:514–518.
84. Feyereisen M, Fitzgerald G, Komornicki A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem Phys Lett.* 1993;208:359–363.
85. Polly R, Werner H-J, Manby FR, Knowles PJ. Fast Hartree–Fock theory using local density-fitting approximations. *Mol Phys.* 2004;102:2311–2321.
86. Jung Y, Lochan RC, Dutoi AD, Head-Gordon M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J Chem Phys.* 2004;121:9793–9802.
87. Fales B, Scott, Curtis Ethan R., Johnson K, Grace, Lahana Dean, Seritan Stefan, Wang Yuanheng, Weir Hayley, Martínez Todd J., Hohenstein Edward G.. Performance of Coupled-Cluster Singles and Doubles on Modern Stream Processing Architectures. *Journal of Chemical Theory and Computation.* 2020;<http://dx.doi.org/10.1021/acs.jctc.0c00336>.
88. Fales BS, Shu Y, Levine BG, Hohenstein EG. Complete active space configuration interaction from state-averaged configuration interaction singles natural orbitals: Analytic first derivatives and derivative coupling vectors. *J Chem Phys.* 2017;147:094104.
89. Snyder JW, Parrish RM, Martínez TJ.  $\alpha$ -CASSCF: An efficient, empirical correction for SA-CASSCF To closely approximate MS-CASPT2 potential energy surfaces. *J Phys Chem Lett.* 2017;8:2432–2437.
90. Fales BS, Seritan S, Settje NF, Levine BG, Koch H, Martínez TJ. Large scale electron correlation calculations: Rank-reduced full configuration interaction. *J Chem Theory Comput.* 2018;14:4139–4150.
91. Peng W-T, Fales BS, Levine BG. Simulating electron dynamics of complex molecules with time-dependent complete active space configuration interaction. *J Chem Theory Comput.* 2018;14:4129–4138.
92. Fales BS, Martínez TJ. Efficient treatment of large active spaces through multi-GPU parallel implementation of direct configuration interaction. *J Chem Theory Comput.* 2020;16:1586–1596.
93. Fales BS, Parrish RM, Martínez TJ. Single and mixed precision direct configuration interaction. 2020 (in preparation).
94. Fales BS, Martínez TJ. Fast transformations between configuration state function and slater determinant bases for configuration interaction. *J Chem Phys.* 2020;152:164111.
95. White S. Density matrix formulation for quantum renormalization groups. *Phys Rev Lett.* 1992;69:2863–2866.
96. Keller S, Dolfi M, Troyer M, Reiher M. An efficient matrix product operator representation of the quantum chemical Hamiltonian. *J Chem Phys.* 2015;143:244118.

97. Pijeu S, Hohenstein EG. Improved complete active space configuration interaction energies with a simple correction from density functional theory. *J Chem Theory Comput.* 2017;13:1130–1146.
98. Filatov M, Liu F, Martínez TJ. Analytical derivatives of the individual state energies in ensemble density functional theory method. I. General formalism. *J Chem Phys.* 2017;147:034113.
99. Liu F, Filatov M, Martínez TJ. Analytical derivatives of the individual state energies in ensemble density functional theory method: II. Implementation on graphical processing units (GPUs). *ChemRxiv.* 2019; chemrxiv.7985657.v1.1–59.
100. Filatov M, Liu F, Kim KS, Martínez TJ. Self-consistent implementation of ensemble density functional theory method for multiple strongly correlated electron pairs. *J Chem Phys.* 2016;145:244104.
101. Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. *Chem Rev.* 2005;105:2999–3093.
102. Lange AW, Herbert JM. A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J Chem Phys.* 2010;133:244111.
103. Liu F, Sanchez DM, Kulik HJ, Martínez TJ. Exploiting graphical processing units to enable quantum chemistry calculation of large solvated molecules with conductor-like polarizable continuum models. *Int J Quantum Chem.* 2019;119:e25760.
104. Lamoureux G, Harder E, Vorobyov IV, Roux B, MacKerell AD. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem Phys Lett.* 2006;418:245–249.
105. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79:926–935.
106. Götz AW, Clark MA, Walker RC. An extensible interface for QM/MM molecular dynamics simulations with AMBER. *J Comput Chem.* 2014;35:95–108.
107. Eastman P, Swails J, Chodera JD, et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol.* 2017;13:e1005659.
108. Sisto A, Glowacki DR, Martinez TJ. Ab initio nonadiabatic dynamics of multichromophore complexes: A scalable graphical-processing-unit-accelerated exciton framework. *Acc Chem Res.* 2014;47:2857–2866.
109. Sisto A, Stross C, Van Der Kamp MW, et al. Atomistic non-adiabatic dynamics of the LH2 complex with a GPU-accelerated: Ab initio exciton model. *Phys Chem Chem Phys.* 2017;19:14924–14936.
110. Li X, Parrish RM, Liu F, Schumacher SILK, Martínez TJ. An ab initio exciton model including charge-transfer excited states. *J Chem Theory Comput.* 2017;13:3493–3504.
111. Goumans TPM, Catlow CRA, Brown WA, Kästner J, Sherwood P. An embedded cluster study of the formation of water on interstellar dust grains. *Phys Chem Chem Phys.* 2009;11:5431–5436.
112. Kästner J, Carr JM, Keal TW, Thiel W, Wander A, Sherwood P. DL-FIND: An open-source geometry optimizer for atomistic simulations. *J Phys Chem A.* 2009;113:11856–11865.
113. Wang L-P, Song C. Geometry optimization made simple with translation and rotation coordinates. *J Chem Phys.* 2016;144:214108.
114. Henkelman G, Jónsson H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J Chem Phys.* 2000;113:9978–9985.
115. Henkelman G, Uberuaga BP, Jónsson H. Climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J Chem Phys.* 2000;113:9901–9904.
116. Weinan E, Ren W, Vanden-Eijnden E. String method for the study of rare events. *Phys Rev B.* 2002;66:523011–523014.
117. Sheppard D, Terrell R, Henkelman G. Optimization methods for finding minimum energy paths. *J Chem Phys.* 2008;128:134106.
118. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys.* 2007;126:014101.
119. Martyna GJ, Klein ML, Tuckerman M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J Chem Phys.* 1992;97:2635–2643.
120. Bussi G, Parrinello M. Accurate sampling using Langevin dynamics. *Phys Rev E.* 2007;75:056707.
121. Niklasson AMN, Tymczak CJ, Challacombe M. Time-reversible Born–Oppenheimer molecular dynamics. *Phys Rev Lett.* 2006;97:123001.
122. Niklasson AMN, Steneteg P, Odell A, et al. Extended Lagrangian Born–Oppenheimer molecular dynamics with dissipation. *J Chem Phys.* 2009;130:214109.
123. Luehr N, Markland TE, Martinez TJ. Multiple time step integrators in ab initio molecular dynamics. *J Chem Phys.* 2014;140:084116.
124. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: New feathers for an old bird. *Comput Phys Commun.* 2014;185:604–613.
125. Glowacki DR, Paci E, Shalashilin DV. Boxed molecular dynamics: A simple and general technique for accelerating rare event kinetics and mapping free energy in large molecular systems. *J Phys Chem B.* 2009;113:16603–16611.
126. Hirai H. Practical hyperdynamics method for systems with large changes in potential energy. *J Chem Phys.* 2014;141:234109.
127. Ong MT, Leiding J, Tao H, Virshup AM, Martínez TJ. First principles dynamics and minimum energy pathways for mechanochemical ring opening of cyclobutene. *J Am Chem Soc.* 2009;131:6377–6379.
128. Luehr N, Jin AGB, Martinez TJ. Ab initio interactive molecular dynamics on graphical processing units (GPUs). *J Chem Theory Comput.* 2015;11:4536–4544.
129. Wang LP, Titov A, McGibbon R, Liu F, Pande VS, Martinez TJ. Discovering chemistry with an ab initio nanoreactor. *Nat Chem.* 2014;6:1044–1048.

130. Santamaria R, Adamowicz L, Rosas-Acevedo H. Microscopic pressure-cooker model for studying molecules in confinement. *Mol Phys*. 2015;113:671–682.
131. Wang LP, McGibbon RT, Pande VS, Martinez TJ. Automated discovery and refinement of reactive molecular dynamics pathways. *J Chem Theory Comput*. 2016;12:638–649.
132. Valleau S, Martínez TJ. Reaction dynamics of cyanohydrins with hydrosulfide in water. *J Phys Chem A*. 2019;123:7210–7217.
133. Meisner J, Zhu X, Martínez TJ. Computational discovery of the origins of life. *ACS Cent Sci*. 2019;5:1493–1495.
134. Das T, Ghule S, Vanka K. Insights into the origin of life: Did it begin from HCN and H<sub>2</sub>O? *ACS Cent Sci*. 2019;5:1532–1540.
135. Ufimtsev IS, Luehr N, Martinez TJ. Charge transfer and polarization in solvated proteins from Ab initio molecular dynamics. *J Phys Chem Lett*. 2011;2:1789–1793.
136. Kulik HJ, Luehr N, Ufimtsev IS, Martinez TJ. Ab initio quantum chemistry for protein structures. *J Phys Chem B*. 2012;116:12501–12509.
137. Zheng M, Reimers JR, Waller MP, Afonine PV. QIR: Quantum-based refinement. *Acta Crystallogr D*. 2017;73:45–52.
138. Xu Y, Waller MP, Zheng M, Afonine PV, Reimers JR, Moriarty NW. Solving the scalability issue in quantum-based refinement: QIR#1. *Acta Crystallogr D*. 2017;73:1020–1028.
139. Ehrlich S, Göller AH, Grimme S. Towards full quantum-mechanics-based protein–ligand binding affinities. *ChemPhysChem*. 2017;18:898–905.
140. Liang R, Liu F, Martínez TJ. Nonadiabatic photodynamics of retinal protonated Schiff base in Channelrhodopsin 2. *J Phys Chem Lett*. 2019;10:2862–2868.
141. Yu JK, Liang R, Liu F, Martínez TJ. First-principles characterization of the elusive fluorescent state and the structural evolution of retinal protonated Schiff base in bacteriorhodopsin. *J Am Chem Soc*. 2019;141:18193–18203.
142. Curchod BFE, Martínez TJ. Ab initio nonadiabatic quantum molecular dynamics. *Chem Rev*. 2018;118:3305–3336.
143. Curchod BFE, Sisto A, Martínez TJ. Ab initio multiple spawning photochemical dynamics of DMABN using GPUs. *J Phys Chem A*. 2017;121:265–276.
144. Hollas D, Šištík L, Hohenstein EG, Martínez TJ, Slaviček P. Nonadiabatic ab initio molecular dynamics with the floating occupation molecular orbital-complete active space configuration interaction method. *J Chem Theory Comput*. 2018;14:339–350.
145. Wolf TJA, Sanchez DM, Yang J, et al. The photochemical ring-opening of 1,3-cyclohexadiene imaged by ultrafast electron diffraction. *Nat Chem*. 2019;11:504–509.
146. Wilkin KJ, Parrish RM, Yang J, et al. Diffractive imaging of dissociation and ground-state dynamics in a complex molecule. *Phys Rev A*. 2019;100:023402.
147. Weir H, Williams M, Parrish RM., Hohenstein EG., Martínez TJ. Nonadiabatic Dynamics of Photoexcited cis-Stilbene Using Ab Initio Multiple Spawning. *The Journal of Physical Chemistry B*. 2020; <http://dx.doi.org/10.1021/acs.jpcc.0c03344>.
148. Asadchev A, Allada V, Felder J, Bode BM, Gordon MS, Windus TL. Uncontracted Rys quadrature implementation of up to G functions on graphical processing units. *J Chem Theory Comput*. 2010;6:696–704.
149. Kussmann J, Ochsenfeld C. Hybrid CPU/GPU integral engine for strong-scaling ab initio methods. *J Chem Theory Comput*. 2017;13:3153–3159.

**How to cite this article:** Seritan S, Bannwarth C, Fales BS, et al. TeraChem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics. *WIREs Comput Mol Sci*. 2021;11:e1494. <https://doi.org/10.1002/wcms.1494>